Plan Title:	**Demographic Surveys Support System**
Plan Number:	**CB-DM-96-01-E**
Plan ID:	**IT**

PART I - INFORMATION TECHNOLOGY ARCHITECTURE PLAN

1. Information Requirements

The Demographic Programs Directorate is responsible for conducting national surveys, producing population estimates and projections, and providing macro- and microdata on population, income, poverty, housing, jobs, and health in the United States and foreign countries. In doing so, it provides statistical survey design, collection, and processing services to other government agencies and private contractors; and provides for content design, analysis, and dissemination of decennial census data. The Demographic area does its best to identify present and potential customers (data users), find out their needs and expectations, and deliver quality data and analysis to help build and evaluate informed public policies.

Incoming survey data are obtained for the most part through the Census Bureau's Field Division, whether through telephone or personal interview. One other major source of data is an administrative records file obtained from the Internal Revenue Service; this data file is used for benchmarking the country's population throughout the decade.

The work performed by Demographic Programs encompasses seven major programs or categories. It includes data analysis as well as survey methodology and data processing functions, which in general are broken down as follows. Mathematical statisticians within the Demographic Statistical Methods Division (DSMD) provide support with tasks such as writing sampling and weighting procedures for the Directorate. Demographic Surveys Division (DSD) is responsible for most of the data processing tasks; these include providing computer-assisted interviewing "instruments" to Field Division for data collection, writing programs for edits, weighting, and tabulation, and some data dissemination functions. Once the final data files have been verified, subject matter experts within Population Division (POP) and Housing and Household Economics Division (HHES) subset, analyze, and further tabulate the data in order to produce reports for release to the general public and for other sponsors. Although the nature of some surveys or processing tasks dictates that responsibilities sometimes cross organizational boundaries, the professionals within each area work together to ensure that the data and reports produced by the Demographic Directorate are timely, accurate, and informative.

A. Current Population Survey (CPS)

The CPS provides monthly estimates of employment, unemployment, and other characteristics of the general labor force, of the population as a whole, and of various subgroups of the population. The survey is sponsored jointly by the Census Bureau and the Bureau of Labor Statistics (BLS). Data from the CPS are used by the BLS to provide the official monthly estimates of employment and unemployment for the nation, as well as reliable estimates for all states and selected metropolitan areas on an average annual basis; and to determine the distribution of funds under the Job Training Partnership Act.

In addition to labor force data, the CPS is used as the primary vehicle to collect supplemental data as requested by other sponsors. The two major examples are the March Annual Demographic Supplement, which is currently the official source of estimates of income and poverty in the United States and also provides information on work experience and migration, and the October Supplement on School Enrollment.

The CPS data are collected through a combination of Computer Assisted Telephone Interviewing (CATI) and Computer Assisted Personal Interviewing (CAPI). Final basic data files and tabulations are produced and transmitted to the BLS for further processing to produce the official labor force estimates.

B. CPS Modernization

Several years ago, the Census Bureau and the BLS recognized that the questionnaire and methodology developed for technology and labor conditions of a generation ago were no longer suitable. They embarked on a modernization effort to improve the measurement of labor force characteristics, to expand the labor force data available, and to improve longitudinal measures.

Two major aspects of the CPS modernization were the redesign of the file generation processing system, and the development of a Census Bureau/BLS shared processing system for data access by both Census and BLS analysts. The new CPS file generation system was implemented in January 1994. Subsequent modernization efforts have focused on the development of the data access system.

Within this new data access system, CPS data and documentation will be shared with BLS. A shared system will better serve analytical needs by allowing interactive access to data as well as longitudinal linking. Both Census Bureau and BLS analysts will have quick and easy access to current CPS data and tabular output, historical CPS data, and CPS documentation, as well as a variety of software tools, through a flexible and sophisticated graphical user interface.

C. Survey of Income and Program Participation (SIPP)

The SIPP is a longitudinal survey which provides an overall measure of the economic well-being of the population as a whole, and profiles those individuals and families who receive benefits from one or more government programs. SIPP collects the source and amount of income received, program participation and eligibility data, and general demographic characteristics to measure the effectiveness of existing Federal, state, and local programs. Data are also used to determine future costs and coverage for programs such as food stamps, and to provide statistics on the distribution of income in the Nation. This survey allows analysts to measure the effects of changing demographics (i.e., divorce, birth of a child, etc.) or labor force status, and to draw conclusions about the effects of changes in government programs.

Through the first quarter of FY 1996, SIPP data were collected using the traditional paper-and-pencil methodology. Beginning in the second quarter, data collection moves to the CAPI approach. Final testing of a new (post-data collection) processing system was completed with the first final production microdata product due by early 1997. The new processing system resides on a UNIX server dedicated to this survey. The last remnants of data processing using the pencil-and-paper methodology will be completed by mid-1997, marking the end of Unisys data processing for SIPP.

D. Sample Redesign

The purpose of this system is to select samples for all demographic surveys from the decennial census files, including the 2000 equivalents of the HEDF, SEDF, ACF, IDF, and STFs. (All of the samples derived from the 1990 census will be exhausted at this time.) These input files will be matched to get sample address files and to assign survey and control codes. The advantages of this system are that storing the sample control files for all demographic surveys in one relational database allows easy cross-referencing and adjustment of the survey samples; helps to provide standardization of data products across surveys; and provides more information about samples through ad hoc queries and analysis.

DSMD personnel are currently working to complete the automation of the 1990 Sample Redesign by automating the preparation of samples for interviewing. The Automated Transfer of Listings and Samples (ATLAS) system will include keying of listing, sample control, and reinterview modules. The keying module is responsible for automating the area sample frames and the Components of Inventory Change (CINCH) portion of the AHS. The sample control module will provide all sample addresses to Field Division via the Master Control System for use in CASES-type interviews; it will also receive address updates from Field Division. The automated reinterview module generates reinterview samples and the CASES reinterview "instrument," as well as providing reinterview data for analysis. Completed modules exist at present on the Demographic ISAR; they, along with modules still under development, will exist in the future on a UNIX workstation.

E. Postcensal and Intercensal Estimates

This system provides updated decennial census statistics, projections, and survey controls which are used to administer federal programs and to allocate approximately $40 billion in Federal funds to state and local jurisdictions. Outputs from this system will also be used by other Demographic Directorate programs to provide estimates of age, sex, race, ethnicity, median household income, per capita income, and poverty rates in various geographic levels. These estimates are provided on a biennial basis. Federal users include the Department of Energy, the Department of Housing and Urban Development, the Department of Health and Human Services, the Department of Agriculture, the Bureau of Labor Statistics, the Federal Election Commission, and Congress. These data are also widely used by state and local governments, the media, and marketers and planners.

Input files used include 1980 and 1990 decennial 100 percent and census sample files, Federal income tax extracts, and data files from other affected federal social programs. A significant input to this system is geocoded microdata obtained from the Internal Revenue Service (IRS). These data, aside from providing income and poverty estimates, are also used to provide migration statistics for state and local jurisdictions, and are an important resource in researching methods to improve future decennial census coverage. Tabulations generated from these files are distributed to the general public by the IRS, and are used for planning and general research purposes.

F. Continuous Measurement

Continuous Measurement (CM) (currently waiting approval) is an alternative method of collecting the traditional long form sample content in the decennial census. The primary goal of CM is to provide annual updates of detailed census characteristics. Each year the CM will provide estimates of the characteristics for all geographic and tabulations units with a population of 250,000 or more such as states, congressional districts, metropolitan areas, counties, and cities; it will also provide accumulated multi-year averages for all small geographic and tabulations areas (less than 250,000 population) for which census data are published. The CM System will enhance and modernize the Federal statistical system by providing reliable small area estimates each year, provide the flexibility to react quickly and efficiently to future data needs, and become a rich source of sample for future surveys.

The CM Survey will be conducted nationwide beginning in January 1999. During the first three years (1999 through 2001), staff will use the Master Address File (MAF) to select 400,000 sample addresses each month. This sample size will allow the Census Bureau to produce estimates for all geographic areas by June 2002, when the decennial census estimates would normally be produced. Beginning in January of 2002, CM will convert from a three-year design to a five-year design. Under the five-year design the sample size will be reduced from 400,000 sample addresses per month to 250,000 sample addresses per month. The survey will continue to produce yearly data for both large and small geographic areas.

G. Reimbursable Programs - surveys falling under this category include:

- American Housing Survey - provides ongoing data on the composition of the Nation's housing at the Federal, state, and local levels for use in determining housing policy and program development. The survey is sponsored by the Department of Housing and Urban Development. The survey has a national sample and a metropolitan sample. The National Survey is conducted biennially and data are collected by personal visit and CATI. The Metropolitan Survey is conducted annually in 41 metropolitan areas interviewed on a rotating basis. Data are collected by personal visit. Output takes the form of both publication tables and public use microdata files.

  A major supplement to the National Survey is the CINCH. Its purpose is to measure the change in the housing inventory between the chosen base year and the current National Survey year.

- American Travel Survey - provides national and state estimates of the number of trips of 100 miles or more one way and includes transportation and facilities used, as well as associated purposes, costs, and length. The survey is sponsored by the Department of Transportation. This survey is conducted via CATI/CAPI methods. This is a one-time survey to be conducted between April 1995 and February 1996. The final data products are scheduled to be delivered by July 1997.

- Current Point of Purchase Survey (CPP) - provides the names and locations of stores at which consumers purchase goods and services. BLS uses these data to select and update outlets included in its Consumer Price Index (CPI) pricing surveys. The survey is conducted annually by personal visit. Beginning in 1997, the survey will be collected using CATI methodology.

- National Survey of College Graduates - provides current estimates of the size and characteristics of the Nation's science and engineering population. The survey is sponsored by the National Science Foundation. The survey is both a cross-sectional and longitudinal study. Together with other data collected by the sponsor, the data serve as the basis for the Scientific and Personnel Data System as mandated by Congress. The survey began in 1993 and is conducted biennially. Data collection is by mail with a follow-up of mail nonrespondents by CATI and personal visit.

- Consumer Expenditure Survey - provides a current and continuous series of data on consumer expenditures and other related characteristics for items of both a durable and nondurable nature—everything from housing and utilities to food and magazine subscriptions. The survey is sponsored by BLS. Data are used in determining the need to revise the Consumer Price Index, and for use in family expenditure studies and other analyses. This is a continuing survey conducted monthly by personal visit and is comprised of two components: the Quarterly Interview Survey and the Diary Survey. The Quarterly Interview Survey obtains data on large expenditures that we expect respondents to recall accurately over a three-month period or for which records are likely to be available. The

Diary Survey provides data on those items not covered in detail in the Quarterly Interview Survey; namely, the small everyday type of purchases as well as small expenditure items that the sample household members purchase occasionally.

- Quick STATS - a standardized survey system to conduct a timely, high quality telephone survey via the Census Bureau's CATI system. The goal of this system is to provide data to the survey sponsor within 45 days of the receipt of survey content.

- 1996 National Survey of Fishing, Hunting, and Wildlife-Associated Recreation - provides current data on fishing, hunting, and wildlife-related activities of a nonconsumptive nature such as feeding, observing, and photographing wildlife. The survey is sponsored by the Fish and Wildlife Service of the Department of Interior. Data outputs will consist of a final data file and national and state reports. Federal and state agencies will use these data to help manage the Nation's fish and wildlife resources. Data collection will be through CATI and CAPI.

- National Crime Victimization Surveys (NCVS) - provides information on crime and those victimized by it. The NCVS is a recurring survey of 11,000 households per month. It has been operating continuously since 1973. Data are gathered for incidents of crime, the percentage of crimes reported to police and injuries, and the economic loss resulting from crime. Extensive annual tabulations and a microdata file are produced for the sponsor, the Bureau of Justice Statistics (BJS) of the Department of Justice. Frequently the base program is supplemented with additional related topics such as the School Crime Supplement.

- National Prisoner Statistics Program (NPS) - NPS is an "umbrella" title for a group of surveys of prisoners and prison populations done for BJS. These include recurring surveys as well as one-time special surveys. The largest of the recurring surveys is the National Corrections Reporting Program (NCRP) which annually collects extensive data on the inmate populations of the 40 participating states. Other recurring surveys which fit under this "umbrella" and which are currently underway include the Survey of Adults on Parole and the Survey of Local Jails - 1995.

- Education Surveys - These surveys are an integrated set of sample and universe surveys conducted for the National Center for Education Statistics (NCES) of the Department of Education. The sample surveys, known collectively as the School and Staffing Surveys or SASS, collect information about elementary and secondary school districts, public, private and Indian schools, teachers, administrators, students, libraries, and librarians. The Private School Survey (PSS) is a universe survey of all private schools having students in grades one through twelve. The Teacher Follow-up Survey (TFS) is done in the year following the SASS surveys to collect information on teachers who have stayed in the profession and those who have left the teaching profession. SASS surveys were conducted in 1988, 1991 and 1993, with the next scheduled for 1998. The PSS is conducted every two years.

- Survey of Program Dynamics (SPD) - SPD was projected as a new longitudinal survey scheduled to begin in 1996. The Department of Health and Human Services was to be the sponsor. Funding for the survey did not materialize; work on this project has been suspended indefinitely.

- Other - in addition, the Demographic Programs Directorate conducts many other reimbursable and periodic surveys on a one-time or recurring basis. These surveys vary in size and complexity. Examples include the New York City Housing Vacancy Survey, the Property Owners and Managers Survey, the Long Term Care Survey, and the National Longitudinal Surveys.

2. Planned Processing and Telecommunications Architecture

Currently, the Demographic Directorate makes use of four processing platforms: Unisys mainframe, DEC VAX/VMS minicomputers, POSIX-compliant workstations, and Novell NetWare PC LANs. The Unisys mainframe is the Census Bureau's legacy computer system and will be decommissioned in late 1997. Consequently, we have plans to migrate - either by re-engineering or converting - all work on that system to another platform. The DEC VAXcluster was originally installed in early 1988 for the purpose of providing tools for end user data analysis; later, some production processing was moved from the Unisys to this machine. POSIX-compliant workstations first came into use in 1990 as an end user replacement for the DEC resources. As their acceptance grew and their performance and flexibility became proven, we acquired more machines, both for expanded data analysis capability and for small survey processing. The Novell network is used almost exclusively for office automation purposes.

Demographic Programs, in accordance with the IT Directorate's open systems decree and the Federal Government's POSIX mandate, has begun to move all of our survey processing, analysis, and dissemination work to UNIX workstations. We chose this type of processing platform for several reasons. First, a distributed workstation environment provides better I/O capability than our large VAX machines. Second, competition among UNIX hardware vendors has made this solution a very attractive price/performance alternative. Third, a distributed processing environment eases the burden of contingency planning while allowing more flexibility and responsiveness than one large centralized platform. Fourth, the availability of uniform off-the-shelf software products for most major brands of UNIX workstations allows us to build and maintain systems which are easily moved between machines. Fifth, a distributed environment provides the capability for shared access with our survey sponsors to intermediate and final survey data and computing resources, a capability for which several sponsors have expressed a desire.

Specifically, we plan to build an integrated network of UNIX workstations and servers to support all of our survey processing, analysis, and dissemination work. We envision a distributed UNIX environment, with one or more servers dedicated to each of our major surveys, programs, and/or activities. We believe this type of environment will provide us with the greatest amount of flexibility to add or reallocate computer resources to meet the different demands and changing

requirements of each program. A distributed processing architecture of this sort will enhance our contingency planning, reducing the risk of interrupting program processing cycles due to hardware or network failure; allows us to meet or exceed our survey obligations and customer expectations by eliminating contention between survey processing cycles; provides more opportunity and flexibility to deal with unexpected requests from our sponsors; and allows for more opportunity to improve our survey development and processing methods. It is our top priority to move all of our current Unisys work to UNIX platforms before the mainframe is decommissioned at the end of 1997.

During FY 1996, the Demographic Directorate established a LAN Steering Committee to oversee both the standardization of all LAN and office automation systems, and to coordinate research into Bureau-wide migration of end user desktop computers to a 32-bit operating system such as Windows NT or Windows 95. We continue to make use of off-the-shelf software in order to ease our evolving migration efforts.

3. Security

A security plan for each sensitive system is on file with the Census Bureau's ADP Security Office. Other plans will be filed as necessary.

Plan Title:	**Demographic Surveys Support System**
Plan Number:	**CB-DM-96-01-E**
Plan ID:	**IT**

PART II - ANNUAL PLAN

1. Architecture Status

The computer hardware used by Demographic Programs is a mixture of a corporate Unisys mainframe, DEC VAX/VMS minicomputers, POSIX-compliant workstations and servers, and Novell NetWare-based PC LANs. Over the next 5 years, that mixture will move steadily away from Unisys and VAX to UNIX workstations and personal computers.

2. IT Objectives

Our IT-related objectives for the coming fiscal year are threefold. First, we will continue our efforts to move all Demographic processing and analysis work off the Unisys mainframe by the end of 1997. Where possible, we will re-engineer our major survey systems to be more flexible and easier to maintain so that we may better meet our customers' expectations. Second, in tandem with the Unisys migration effort, we will continue to move our processing and analysis work to an open systems (UNIX) computer environment. Our progress in this area is entirely dependent upon our ability to acquire UNIX resources. Included in this migration process is a concerted effort to reduce our dependence upon non-POSIX network protocols such as DEC's LAT. Our third goal is to reduce the number and size of our current paper reports and instead to place more emphasis on data dissemination through electronic means such as CD-ROM and Internet access.

3. Status

A. Current Population Survey (CPS)

Processing of the CPS continued on the DEC VAX platform. We continue to transmit data to BLS through the Census/BLS shared processing system. We plan to migrate CPS production processing to a UNIX environment to more fully integrate it with the shared processing system and BLS production processing. We have acquired UNIX resources, and DSD staff have begun work on this effort.

1996 was a bridge year for the March CPS processing. Production processing was done as usual on the Unisys mainframe. An internal-use-only file was delivered in July. Programs and data will be moved to the CPS UNIX development platform. The existing I/O system will be replaced by a relational database with an SQL interface and the data will be reprocessed. Once completed, work will begin in HHES to re-engineer their production processing of the data; this work should be completed in late 1997.

In 1997, review of these parallel processed files will continue, leading to production processing on the UNIX platform in April 1997. Tabulations will continue to be processed on the Unisys mainframe until a set of revised specifications is completed in 1997. Our ultimate goal is to move all processing to the UNIX environment.

Milestones	Date of completion (Month-Year)
Functional requirements	
Base CPS	Completed
March CPS (DSD)	August 1995
March CPS (HHES)	May 1997
Platform selection	
Base CPS	Completed
March CPS (DSD)	Completed
March CPS (HHES)	May 1996
Procurement	
Base CPS - Development platform	Completed
Base CPS - Production platform	December 1996
March CPS supplement (DSD)	Completed
March CPS supplement (HHES)	June 1996
System development	
Base CPS UNIX migration	June 1997
March CPS supplement (DSD)	March 1997 (DSD)
March CPS supplement (HHES)	December 1997 (HHES)
Implementation	
Base CPS UNIX migration	July 1997
March CPS supplement (DSD)	March 1997
March CPS supplement (HHES)	December 1997

B. CPS Modernization

The development of the Census/BLS shared processing system continued. CPS basic and supplemental microdata from January 1994 forward is available to Census and BLS analysts. We developed a prototype Internet graphical user interface tool, named FERRET, to allow analysts to extract and analyze CPS basic data from the database on an ad hoc basis and link the data to electronic documentation. We released a Census/BLS supplement Internet site with March 1995 supplement data as the first data set and with FERRET as the data access tool.

We hope to accomplish several objectives next year: 1) transfer more of BLS' production processing to the shared system; 2) make CPS macrodata available in the form of tables and time series; 3) extend electronic documentation and strengthen its link to the data, so that electronic access of data will provide access to thorough documentation, and electronic access to documentation will provide faster and more flexible access to the data; 4) improve database access performance; 5) generalize the system to handle surveys other than CPS; 6) provide an electronic documentation facility; and 7) begin development of the production version of FERRET.

Milestones		**Date of completion (Month-Year)**
Functional requirements		Completed
Platform selection		Completed
Procurement		Ongoing
System Development/Implementation (Phased prototypes)		
Phase 1 -	No GUI; microdata available through SQL query access only	Completed
Phase 2 -	Limited GUI; interface allows users to extract data, issue queries, and produce macrodata; limited documentation and functionality	Prototype -completed Production - January 1997
Phase 3 -	Sophisticated user interface; will allow users to explore whatever data, documentation, and processes are available; access to full documentation; linking of data and documentation	January 1998
Phase 4 -	Access to historical CPS data	January 1999
Phase 5 -	Remaining functional requirements	January 2001

C. Survey of Income and Program Participation (SIPP)

In 1996, we:
- nearly completed data processing for the last paper-and-pencil version of SIPP. Remaining tasks will be completed by mid-1997.
- completed the processing and analysis of dress rehearsal data collected using the new CAPI technology.
- began the processing of the new SIPP panel collected using CAPI.

In 1997, we will:
- complete all data processing of data collected using pencil and paper.
- complete the processing of core data collected in the first three sets of interviews.
- develop additional modules for processing topical data collected in interviews three and four.
- develop a multi-wave or longitudinal processing system to combine the wave data into a more usable format.
- make the first version of the user interface system available.

Milestones	Date of completion (Month-Year)
Functional requirements	
Base SIPP (DSD)	Completed
HHES production table processing	December 1996
Platform selection	
Base SIPP (DSD)	Completed
HHES production table processing	February 1997
Procurement	
Base SIPP - development platform	Completed
Base SIPP - production platform	Completed
HHES production table processing	January 1998
Software development/testing	
Cross-sectional Wave 1	January 1996
Cross-sectional Wave 2	July 1996
Longitudinal	August 1997
HHES production table processing	June 1998

Milestones	Date of completion (Month-Year)
Implementation	
Cross-sectional Wave 1	April 1996
Cross-sectional Wave 2	August 1996
Longitudinal	September 1997
HHES production table processing	September 1998

D. Sample Redesign

DSMD personnel continued the phase-in of the 1990 design sample for the current surveys with the introduction of the new samples for CE and SIPP. They performed sample adjustments to NHIS, CPS, and SIPP to respond to changing sponsor requirements and budget conditions. They completed the phase-out of the 1980 sample-based processing on the Unisys mainframe. Remaining Unisys work is limited to converting historical data, and the New Construction PAL system. The PAL system is itself ready for migration to an open environment; this will occur when the CASIC CAPI technologies for automated listing of PAL and SOC data are complete.

DSMD programmers and analysts spent extensive time in the specification and design of the ATLAS components for Listing and Sample Control. A test data base was developed in Oracle on a UNIX server to provide a base for system development and testing. Analysts and programmers continue to develop plans specifications to convert CINCH from paper to an automated system.

The generic Reinterview system was completed and put into production for five surveys: ATS-RE, ATS-QC, NSCG, NLS, and SIPP dress rehearsal. Work continues on building better interfaces and production controls with TMO and CASIC, investigating the implications of using the next version of CASES software, and preparing to put additional surveys into the field.

Milestones	Date of completion (Month-Year)
Functional requirements - ATLAS	Completed
Platform selection - ATLAS	Completed
Procurement - ATLAS	Completed
Software development/testing - ATLAS	June 1997
Implementation - ATLAS	January 1998
Functional requirements - Sample redesign	January 2000

Milestones	Date of completion (Month-Year)
Platform selection - Sample redesign	January 2000
Procurement - Sample redesign	June 2000
Software development/testing - Sample redesign	December 2003
Implementation - Sample redesign	April 2004

E. Postcensal and Intercensal Estimates

Population Division personnel are well on their way in the effort to move this processing system from the Bureau's Unisys mainframe to a UNIX server. The new system will be operational in July of 1996, and most Unisys processing will be converted by December. Because of the large number of IBM 3480 cartridges which must be converted, Population Division's final migration from the Unisys will run into 1997.

Outputs from the Population system will feed into HHES's Small Area Income and Poverty Estimates program, which provides income and poverty statistics down to the county level. This program is in place. Software for producing the 1993 and 1995 estimates will finish the development and testing stage in March 1996 and March 1998, respectively. This system will also provide input data for the National and State Projections program, which produces population projections by age, sex, and ethnicity.

Milestones	Date of completion (Month-Year)
Functional requirements	Completed
Platform selection	
SAIPE (HHES)	Completed
Production processing (POP)	Completed
Procurement	
SAIPE (HHES)	Completed
Production processing (POP)	June 1996
Software development/testing	
1993 SAIPE (HHES)	March 1996
Production processing (POP)	August 1996
1995 SAIPE (HHES)	March 1998

Milestones	Date of completion (Month-Year)
Implementation	
1993 SAIPE (HHES)	March 1996
Production processing (POP)	October 1996
1995 SAIPE (HHES)	March 1998

F. Continuous Measurement

The 1996 CM Test will run from November 1995 through December 1996 and will encompass 7,500 sample units in four test sites, along with 4000 sample units in a national sample component. Interviewing will test a combination of self-enumeration, CATI, and CAPI techniques. In FY 1997 we will continue testing using a much larger sample. At this time, we still intend to begin production in 1999.

The CM Office officially became a part of Demographic Statistical Methods Division in February 1996.

Milestones	Date of completion (Month-Year)
Functional requirements	
DSMD	December 1996
CM Staff	December 1996
Platform selection	
DSMD	December 1996
CM Staff	June 1997
Procurement	
CM Staff	March 1998
Software development/testing	
DSMD	June 1997
CM Staff	December 1998
Implementation	
DSMD	September 1998
CM Staff	January 1999

G. Reimbursable projects

American Housing Survey (AHS)

The AHS processing system currently resides on the Unisys mainframe. The 1995 and 1996 surveys will use the existing processing system, while the AHS is undergoing a complete re-engineering. The 1997 surveys will include a new questionnaire, computer-assisted data collection, and a redesigned processing system. Our target platform will be a UNIX server. In re-engineering the AHS, we hope to meet the sponsor's goals of improved data quality, more timely delivery of data, and improved responsiveness.

We have seven process teams developing the functional requirements for the new processing system. All of the teams except one should complete their work by mid-March 1996. Immediately thereafter, another team will begin the design of the new processing system. At that point implementation will begin in stages. The pre-enumeration processing system must be in place by the end of FY 1996.

Milestones	Date of completion (Month-Year)
Off Unisys 1995 National 1996 MS 1993 CINCH	 December 1996 December 1997 March 1997
Functional requirements - Redesign	December 1995
Platform selection	January 1996
Procurement Development platform Production platform Census/HUD shared data access platform	 February 1996 December 1996 October 1997
Software development/testing	November 1998 (flow basis)
Implementation	September 1996 - December 1998 (flow basis)

Current Point of Purchase Survey (CPP)

In FY 1996 we completed the processing of the 1995 Paper and Pencil survey, and are preparing to process the 1996 survey. This system resides on the Unisys mainframe; 1996 will be the last year the survey is collected via paper and pencil and processed on the mainframe. In FY 1996 we will continue to process a test version of the CPP using CATI data collection methodologies. This processing system, which also includes an interactive coding system, resides on the Demographic ISAR VAXcluster and uses RDB, Fortran, and SAS. The system will be the basis for the eventual production system that will be used in 1997. We hope to begin migration of the

CPP CATI processing system from the VAX to a UNIX server in FY 1996 prior to production implementation in 1997, pending acquisition of a UNIX platform.

Milestones	Date of completion (Month-Year)
Off Unisys - 1996 CPP	December 1996
Functional requirements - 1997 CPP CATI (CPP Redesign)	September 1996
Platform selection	November 1996
Procurement	December 1996
Migration to UNIX	December 1996
Implementation	January 1997

National Survey of College Graduates (NSCG)

In FY 1996 we completed 1995 NSCG data collection activities. We have begun data processing and expect to release a user file to the sponsor in August.

In FY 1997, we will select a sample of respondents for the 1997 survey and begin data collection activities. We will then process the mail returns and identify nonrespondents for follow-up interviewing.

Milestones	Date of completion (Month-Year)
Functional requirements	N/A
Platform selection	October 1995
Procurement	November 1995
System Development	August 1996
Implementation	January 1995 - August 1996 (flow basis)

Consumer Expenditure Survey (CES)

The current CES processing system resides on the Unisys mainframe, although migration efforts are underway. The Census Bureau worked with the sponsor, the BLS, to develop requirements for a redesign of this processing system. These requirements were completed in January 1995. At that point, the BLS decided that they would assume responsibility for the design, development, and implementation of the post collection processing (known as the Redesign Phase Two System). Census would maintain responsibility for data collection, data keying, and preprocessing

activities (known as the Redesign Phase One System). Our plans call for continuing processing on the Unisys mainframe with the current system through FY 1997.

The BLS has numerous requirements for data receipt in FY 1997. They want to receive the data in the traditional format, as well as the new format which they will receive once they assume responsibility for processing the data. We will also implement a data transfer system during FY 1997 similar to that used for the Current Population Survey.

Milestones	Date of completion (Month-Year)
Functional requirements (data transfer)	October 1996
Platform selection	December 1996
Procurement	January 1997
System Development	August 1997
Implementation	October 1997

National Crime Victimization Survey (NCVS)

In fiscal year 1996 all production processing to create microdata files was done on the Unisys mainframe. Tabulations were created using SAS on DEC minicomputers. The development of functional requirements for a new system will be completed this fiscal year.

In fiscal 1997 the goal is to procure an open system to replace the functionality of the Unisys mainframe while developing a SAS-based processing system on an existing UNIX workstation.

Milestones	Date of completion (Month-Year)
Functional requirements	June 1996
Platform selection	January 1997
Procurement	March 1997
System Development	March 1997
Implementation	January 1998

National Prisoner Statistics Program (NPS)

In 1996 all processing for NPS surveys was conducted on a UNIX workstation using SAS software. No migration is required. In fiscal 1997 we intend to continue processing as in 1996, with a possible replacement/upgrade of hardware planned for late FY 1997.

Milestones	Date of completion (Month-Year)
Functional requirements (replacement/upgrade)	January 1997
Platform selection	February 1997
Procurement	March 1997
System Development	N/A
Implementation	N/A

Education Surveys (SASS, TFS, PSS)

In fiscal 1996, all processing for the 1994/1995 TFS and the 1995/1996 PSS surveys was completed using SAS software on a UNIX workstation. Tabulations and tape creation were done using SAS on DEC minicomputers and office microcomputers. No major education surveys will be processed during FY 1997. However, research, special requests, and support services using past education data sets will be performed. Also, we will be doing preparatory work for the 1997/1998 PSS, and possibly processing and pre-tests for the 1998 SASS surveys. No further migration is expected.

Milestones	Date of completion (Month-Year)
Functional requirements (replacement/upgrade)	October 1997
Platform selection	January 1998
Procurement	June 1998
System Development	June 1999
Implementation	July 1999

Other Reimbursable Programs

- American Travel Survey

We delivered a preliminary weighted file to the sponsor. CASES was used for the survey data collection system. We plan to deliver final edited and weighted data files and tables to the sponsor in fiscal year 1997. The survey is processed on the DEC VAX/VMS and UNIX platforms.

- NHIS (National Health Interview Survey)

  Sponsored by NCHS, DSMD processes an additional permit sample each month. On a quarterly basis, we provide sample controls and addresses to TMO and a weighting control to NCHS.

- 1996 National Survey of Fishing, Hunting, and Wildlife-Associated Recreation

  We began initial planning for this survey in FY 1995. Data collection is scheduled to begin in FY 1996 as well as system development. The survey will be processed in a UNIX environment. Data products will be delivered by February 1997.

- Other

  These include other surveys conducted on a one-time or periodic basis, and vary in size and complexity. Examples include the New York City Housing Vacancy Survey, the Property Owners and Managers Survey, the Long Term Care Survey, the Residential Finance Survey, the Survey of Market Absorption, and the National Longitudinal Surveys. The surveys are processed primarily in SAS on either the DEC VAX/VMS or UNIX platforms, depending on availability of UNIX computing resources. Our goal is to move completely into the UNIX environment, pending acquisition of additional UNIX resources.

CAI Instrument Programming and the Quick Survey Turnaround and Tabulation System (Quick STATS)

CAI instrument programming was completed for a number of surveys which are moving to CATI/CAPI for the first time. In 1996, we launched production versions of SIPP Wave 1 and Wave 2 instruments. We also began production for the National Survey of Fishing, Hunting, and Wildlife-Associated Recreation, the National Survey of Homeless Assistance Providers and Clients, and a production test of the National Health Interview Survey (NHIS). Still other surveys have been updated or enhanced for 1996. We have also completed the software portion of a system which automates the receipt of survey data from the Master Control System operated by the Technology Management Office (TMO). We expect to production test our QuickSTATS system later this year with a survey for NCHS.

In 1997, we will deliver production versions of the NHIS, the American Housing Survey, and others. We will acquire a UNIX workstation to provide a platform for retrieval of survey data,

provide additional required storage, and allow for the proper backup of critical data. The workstation will also provide program development capabilities for new CAI systems.

Milestones	Date of completion (Month-Year)
Functional requirements	July 1997
Platform selection	August 1997
Procurement	October 1997
System Development	November 1997
Implementation	November 1997

General support infrastructure

- PC, LAN, and UNIX

In fiscal year 1996, DSD completed the conversion to new Novell file servers and a new file backup system, and continued replacement of PCs based on a 4-year life cycle. We also completed the first phase of migration to a Windows environment, consisting of installation of Windows-based software with continued access to DOS-based programs. The second phase, to be completed in FY 1997, will consist of upgrading all DOS-based software to Windows-based software, providing necessary training for DSD staff, and providing a total Windows environment.

Also in FY 1996, DSD incorporated additional UNIX systems from other DSD branches under centralized DSD support and acquired a dedicated systems development/testing/certification machine. Also, we upgraded our file backup system, including provisions for secure subnets and functionally separate projects and surveys; and installed a UNIX print server for network-based printing to our existing multi-protocol printers. In FY 1997, we hope to establish electronic documentation and inventory of our UNIX systems, perform research into and acquire automated system management tools, acquire a help desk system to monitor service requests, and strengthen our user and storage management policies.

DSMD nearly completed its migration to a Microsoft Windows-based PC/LAN environment. Most software packages were upgraded to Windows versions, and all DSMD employees attended a Windows training class taught by DSMD LAN support personnel. In addition, DSMD acquired two new UNIX workstations to support database and SAS applications, and upgraded the file backup system to incorporate robotic hardware. In FY 1997, DSMD plans to upgrade their Novell file servers; upgrade all other PC/LAN software packages to Windows versions; replace all 80386 personal computers with Pentiums; and upgrade an older UNIX system.

The CM Office completed the installation of Windows-based networked personal computers using FTP Inc.'s PC/TCP software. In FY 1997, CM will begin using OnTime software for scheduling.

HHES and Population Divisions have nearly completed the replacement of all dumb terminals with networked personal computers for division staff. This will allow them to using cc:Mail and OnTime, the Bureau's standard E-mail and electronic calendaring packages, as well as take advantage of the graphical features offered by most modern software packages.

- Generalized Software Development and Support

A small group of DSD programmers provides troubleshooting and generalized software support to the rest of the division. In 1996, this group controlled DEC disc allocations and balancing, monitored division use of Unisys mass storage, provided technical support to other branches as required, and developed the following software for division use:

1. A generalized SAS based software system to simplify and standardize the production of complex tabulations (DSDTABS). Extensions to DSDTABS, allowing publication-quality table output and user-friendly (GUI) specification of tables, are under development.
2. A generalized system to simplify the use of data collected with CASES software. (CASES Extract System)

In 1997 this group is expected to conclude support for applications running on the Unisys mainframe. Support activities for the DEC minicomputers and UNIX workstations will continue and grow. As we migrate from the Unisys and DEC toward a UNIX environment, their efforts will shift as well to UNIX. Their focus, however, will be on generalized software support rather than system support.